

Problem Statement: Customer Churn in Telecommunications

Telecommunications companies face challenges in retaining customers due to factors such as service quality, pricing, and competition. Customer churn, the rate at which customers switch from one provider to another, is a critical metric for business success. Identifying customers at risk of churn and implementing targeted retention strategies is crucial for reducing revenue loss and maintaining market share.

The objective of this project is to develop a predictive model that can accurately identify customers likely to churn based on their demographic and service usage data. By proactively identifying potential churners, telecommunications companies can implement retention initiatives to mitigate churn and improve customer satisfaction.

Dataset Overview

The dataset consists of 7043 observations and 20 features, including:

- Demographic Information: Gender, SeniorCitizen, Partner, Dependents.
- Service Details: Tenure, PhoneService, MultipleLines, InternetService, OnlineSecurity, OnlineBackup,
 DeviceProtection, TechSupport, StreamingTV, StreamingMovies.
- Contract and Billing: Contract, PaperlessBilling, PaymentMethod, MonthlyCharges, TotalCharges.
- Target Variable: Churn (Whether the customer churned or not).

Data Preprocessing

Handling Missing Values: There were no missing values in the dataset.

Data Cleaning: Removed rows with missing values in the 'TotalCharges' column.

Feature Engineering: Converted column names to lowercase for consistency. Dropped the 'customerID' column as it was not relevant for modeling.

Exploratory Data Analysis (EDA): Conducted EDA to understand the distribution of features and their relationship with the target variable. Visualized features such as 'TotalCharges', 'Tenure', and 'MonthlyCharges' using box plots to identify outliers.

Feature Encoding: Encoded categorical features using LabelEncoder, converting categorical variables into numerical representations.

Model Selection and Evaluation

Evaluated each model's performance using 5-fold cross-validation:

- Random Forest: Train Score: 0.7869, Test Score: 0.7765
- **Decision Trees**: Train Score: 0.7867, Test Score: 0.7765
- Logistic Regression: Train Score: 0.7850, Test Score: 0.7770
- K-Nearest Neighbors: Train Score: 0.7723, Test Score: 0.7651
- Naive Bayes: Train Score: 0.7361, Test Score: 0.7287

Model Deployment

Selected the Random Forest Classifier as the final model and deployed it using pickle serialization.



Key Findings and Insights

Demographic Trends

Analysis revealed notable demographic trends impacting churn rates, such as age, location, and income levels.

2 Usage Patterns

Insights into customer behavior and usage patterns were obtained, shedding light on factors influencing churn decisions.

3 Subscription Plans

Findings regarding the correlation between different subscription plans and churn rates were identified, providing valuable insights.

Factors Contributing to Customer Churn

Customer Service

Addressing the impact of customer service experiences on churn rates and customer satisfaction.

Competitor Analysis

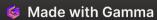
Evaluating the influence of competitor offerings and market dynamics on customer attrition.

Product Performance

Assessing the relationship between product quality, performance, and customer retention.

Pricing Strategies

Exploring the role of pricing and promotional strategies in customer churn dynamics.



Conclusion

In conclusion, the Random Forest model, with its superior predictive accuracy, offers the telecommunications company a robust tool to identify customers at risk of churning. By leveraging insights from this model, the company can strategically allocate resources towards retention efforts, thereby maximizing customer lifetime value and fostering sustainable growth. Continued refinement and feature engineering will enable the company to enhance the model's effectiveness in predicting churn and implementing targeted retention strategies, ultimately strengthening customer loyalty and competitiveness in the market.

